

WHAT IS CLAIMED IS:

1. A method for speech recognition, comprising the steps of:

generating a set of likely hypotheses in recognizing
5 speech;

rescoring the likely hypotheses by using semantic content by employing semantic structured language models; and

scoring parse trees to identify a best sentence
10 according to the sentences' parse tree by employing the semantic structured language models to clarify the recognized speech.

2. The method as recited in claim 1, further
15 comprising the step of training a language model using speech recognition methods.

3. The method as recited in claim 1, wherein the set of likely hypotheses is in the form of an N-best list or
20 lattice.

4. The method as recited in claim 1, wherein the step of rescoring employs MELM2 or MELM3 semantic structured language models.

5. The method as recited in claim 1, wherein the step of scoring parse trees to identify a best sentence according to the sentence's parse tree by employing the semantic structured language models to clarify the recognized speech includes the step of training the structured semantic language models in accordance with history parameters and history questions.

6. The method as recited in claim 5, wherein the history parameters include a previous word (wj-1), a previous word of the previous word (wj-2), a parent constituent label (L), a number of tokens (N) to the left since L starts, a previous closed constituent label (O), a number of tokens (M) to the left after O finishes, and a grandparent label (G).

7. The method as recited in claim 5, wherein the history questions include a default, (wj-1), (wj-1, wj-2), (L,N), (O,M), and (L,G).

8. The method as recited in claim 1, further comprising the step of determining a confidence measurement.

5 9. The method as recited in claim 8, wherein the step of determining a confidence measurement includes employing a statistical method to combine word sequences with a parser tree to determine a confidence score for recognized speech.

10

10. The method as recited in claim 8, wherein the step of determining a confidence measurement includes employing scores obtained from the semantic structured language models along with other speech recognition based
15 features.

11. The method as recited in claim 1, further comprising the step of extracting probabilities assigned to tags, labels and extensions obtained from a parser tree.

20

12. The method as recited in claim 11, further comprising the step of combining the semantic structured

language models and speech recognition based features with the extracted probabilities using a classifier.

13. The method as recited in claim 1, wherein the
5 semantic structured language models are trained by employing unigram, bigram and trigram features.

14. The method as recited in claim 1, wherein the
semantic structured language models are trained using one
10 or more of relative labels, token numbers, and answers to questions related to word order or placement.

15. The method as recited in claim 1, wherein the
semantic structured language models are trained by
15 including a unigram feature, a bigram feature, a trigram feature, a current active parent label (L_i), a number of tokens (N_i) to the left since current parent label (L_i) starts, a previous closed constituent label (O_i), a number of tokens (M_i) to the left after the previous closed
20 constituent label finishes, and a number of questions to classify parser tree entries.

16. The method as recited in claim 15, wherein the questions include a default, (wj-1), (wj-1, wj-2), (Li), (Li, Ni), (Li, Ni, wj-1), and (Oi, Mi), where w represents a word and j is an index representing word position.

5

17. A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for speech recognition, in accordance with claim 1.

10

18. A method for speech recognition, comprising the steps of:

generating a set of likely hypotheses in recognizing speech;

15

rescoring the likely hypotheses by using semantic content by employing semantic structured language models; and

20

scoring parse trees to identify a best sentence according to the sentence's parse tree by employing the semantic structured language models to clarify the recognized speech; and

determining a confidence measurement by employing scores obtained from the semantic structured language models along with other speech recognition based features.

5 19. The method as recited in claim 18, wherein the set of likely hypotheses is in the form of an N-best list or lattice.

10 20. The method as recited in claim 18, wherein the step of rescoring employs MELM2 or MELM3 semantic structured language models.

15 21. The method as recited in claim 18, wherein the step of scoring parse trees to identify a best sentence according to the sentence's parse tree by employing the semantic structured language models to clarify the recognized speech includes the step of training the semantic structured language models in accordance with history parameters and history questions.

20

22. The method as recited in claim 21, wherein the history parameters include a previous word (wj-1), a previous word of the previous word (wj-2), a parent

constituent label (L), a number of tokens (N) to the left since L starts, a previous closed constituent label (O), a number of tokens (M) to the left after O finishes, and a grandparent label (G).

5

23. The method as recited in claim 21, wherein the history questions include a default, (wj-1), (wj-1, wj-2), (L,N), (O,M), and (L,G).

10

24. The method as recited in claim 18, further comprising the step of extracting probabilities assigned to tags, labels and extensions obtained from a parser tree.

15

25. The method as recited in claim 24, further comprising the step of combining the semantic structured language models and speech recognition based features with the extracted probabilities using a classifier.

20

26. The method as recited in claim 18, wherein the semantic structured language models are trained by employing unigram, bigram and trigram features.

27. The method as recited in claim 18, wherein the semantic structured language models are trained using one or more of relative labels, token numbers, and answers to questions related to word order or placement.

5

28. The method as recited in claim 18, wherein the semantic structured language models are trained by including a unigram feature, a bigram feature, a trigram feature, a current active parent label (L_i), a number of tokens (N_i) to the left since current parent label (L_i) starts, a previous closed constituent label (O_i), a number of tokens (M_i) to the left after the previous closed constituent label finishes, and a number of questions to classify parser tree entries.

10

15

29. The method as recited in claim 28, wherein the questions include a default, (w_{j-1}), (w_{j-1} , w_{j-2}), (L_i), (L_i , N_i), (L_i , N_i , w_{j-1}), and (O_i , M_i), where w represents a word and j is an index representing word position.

20

30. A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for speech recognition,

in accordance with claim 18.

31. A system for speech recognition, comprising:

a unified language model including a semantic language
5 model and a lexical language model;

a recognition engine which finds a parse tree to
analyze a word group using the lexical model and the
semantic models, wherein the parse tree is selected based
on lexical information and semantic information which
10 considers tags, labels, and extensions to recognize speech.

32. The system as recited in claim 31, wherein the
parser tree includes semantic information and classer
information used in identifying a best parser tree for a
15 given word group.

33. The system as recited in claim 31, wherein the
parser tree includes information extracted from parsed
sentences to statistically model semantic and lexical
20 content of sentences.

34. The system as recited in claim 31, wherein the semantic language model includes unigram, bigram and trigram features.

5 35. The system as recited in claim 31, wherein the semantic language model includes one or more of relative labels, token numbers, and answers to questions related to word order or placement.

10 36. The system as recited in claim 31, wherein the semantic model is trained by including a unigram feature, a bigram feature, a trigram feature, a current active parent label (L_i), a number of tokens (N_i) to the left since current parent label (L_i) starts, a previous closed constituent label (O_i), a number of tokens (M_i) to the left
15 after the previous closed constituent label finishes, and a number of questions to classify parse tree entries.

20 37. The system as recited in claim 36, wherein the questions include a default, (w_{j-1}), (w_{j-1} , w_{j-2}), (L_i), (L_i , N_i), (L_i , N_i , w_{j-1}), and (O_i , M_i), where w represents a word and j is an index representing word position.

38. The system as recited in claim 31, wherein the semantic model is trained by including history parameters and history questions.

5 39. The system as recited in claim 38 wherein the history parameters include a previous word (wj-1), a previous word of the previous word (wj-2), a parent constituent label (L), a number of tokens (N) to the left since L starts, a previous closed constituent label (O), a
10 number of tokens (M) to the left after O finishes, and a grandparent label (G).

40. The system as recited in claim 39, wherein the history questions include a default, (wj-1), (wj-1, wj-2),
15 (L,N), (O,M), and (L,G).

41. The system as recited in claim 31, further comprising a confidence measurement module.

20 42. The system as recited in claim 31, wherein the confidence measurement module employs a statistical method to combine word sequences with the parse tree to determine a confidence score for recognized speech.

43. The system as recited in claim 31, wherein the confidence measurement module extracts probabilities assigned to tag nodes, label nodes and extensions in the
5 parse tree.